The global impact of Wolbachia on mitochondrial diversity and evolution

Marie Cariou, L. Duret, S. Charlat

The global impact of *Wolbachia* on mitochondrial diversity and evolution

Marie Cariou[1,2], Laurent Duret[1] & Sylvain Charlat[1*].

1 : Université de Lyon ; Université Lyon 1 ; CNRS ; UMR 5558, Laboratoire de Biométrie et Biologie Evolutive, 43 boulevard du 11 novembre 1918, Villeurbanne F-69622, France

2 : Current address: University of Namur, Department of Biology, Rue de Bruxelles 61, 5000 Namur, Belgium

*Corresponding author. Email: sylvain.charlat@univ-lyon1.fr - Phone : 33 4 72 44 81 01 - Fax: +33 4 72 43 13 88

Keywords: Wolbachia, mitochondria, genetic diversity, effective population size, selective sweep

Running title: *Wolbachia* and mitochondrial evolution

Abstract

The spread of maternally inherited microorganisms, such as *Wolbachia* bacteria, can induce indirect selective sweeps on host mitochondria, to which they are linked within the cytoplasm. The resulting reduction in effective population size might lead to smaller mitochondrial diversity and reduced efficiency of natural selection. While documented in several host species, it is currently unclear if such a scenario is common enough to globally impact the diversity and evolution of mitochondria in *Wolbachia*-infected lineages. Here we address this question using a mapping of *Wolbachia* acquisition / extinction events on a large mitochondrial DNA tree, including over 1,000 species. Our analyses indicate that on a large phylogenetic scale, other sources of variation, such as mutation rates, tend to hide the effects of *Wolbachia*. However, paired comparisons between closely related infected and uninfected taxa reveal that *Wolbachia* is associated with a twofold reduction in silent mitochondrial polymorphism, and a 13% increase in non-synonymous substitution rates. These findings validate the conjecture that the widespread distribution of *Wolbachia* infections throughout arthropods impacts the effective population size of mitochondria. These effects might in part explain the disconnection between genetic diversity and demographic population size in mitochondria, and also fuel red-queen-like cytonuclear coevolution through the fixation of deleterious mitochondrial alleles.

Introduction

Variations in population size have deep consequences on molecular evolution: small populations harbor fewer polymorphic sites and accumulate deleterious mutations at faster rates because of the predominance of drift. However, other, non-demographic processes, such as intense episodes of selection, also affect genetic diversity and substitution rates, which tends to uncouple the true population size from its abstract genetic counterpart, the effective population size (Ne). Mitochondrial DNA (mtDNA), although commonly used as a genetic marker for a number of good reasons (notably, technical ease-of-use and high mutation rates) is notoriously subject to such disconnection between the true and effective population size (Hurst & Jiggins, 2005; Bazin *et al.*, 2006; Galtier *et al.*, 2009). Here we test the hypothesis that *Wolbachia* bacteria might be part of the explanation, by investigating their global effects on patterns of mitochondrial diversity and evolution.

These intracellular (and thus maternally inherited) symbionts display an impressive variety of effects that make them invasive (O'Neill *et al.*, 1997; Werren *et al.*, 2008; Martinez *et al.*, 2014). They can kill male embryos or turn them to females, reallocating part or all of the reproductive efforts toward the transmitting sex. They can also impede the reproduction of uninfected females, using infected males as sterilizing weapons. *Wolbachia* also commonly provides protection against natural enemies such as viruses, and thus indiscriminately benefits individuals of both sexes. In any case, if the net fitness gain to the infected maternal lineage is sufficient, *Wolbachia* can increase in frequency and drag along the mitochondrial lineage with which it happens to be associated, because the two are genetically linked through maternal transmission (Turelli *et al.*, 1992).

If *Wolbachia* is perfectly transmitted from mothers to offspring, this process ends with the fixation of both the symbiont and the associated mitochondria, erasing the pre-existing

molecular diversity. If transmission is imperfect, *Wolbachia* does not get fixed, but reaches an equilibrium prevalence, where selection for the infected lineage is balanced by imperfect transmission (O'Neill *et al.*, 1997). Interestingly, even in that case, the ancestral mitochondrial polymorphism is erased in the long run, because all uninfected lineages ultimately originate from infected mothers, through imperfect transmission (Turelli *et al.*, 1992). Thus, depending on the stage of the infection, the reduction in polymorphism within a species should either affect only its infected portion, or also extend to the uninfected individuals if the transmission / selection balance has been reached.

A number of case studies have demonstrated that the spread of *Wolbachia* can indeed affect the mtDNA polymorphism (Turelli *et al.*, 1992; Solignac *et al.*, 1994; Ballard *et al.*, 1996; Keeling *et al.*, 2003; Charlat *et al.*, 2009; Graham & Wilson, 2012; Richardson *et al.*, 2012; Schuler *et al.*, 2016). Shoemaker (2004) also provided evidence for an elevated non-synonymous substitution rate in an infected *Drosophila* species compared to its uninfected sister species, making *Wolbachia* and reduction in Ne a very plausible explanation.

While these studies indicate that *Wolbachia* can have an effect on mitochondrial diversity and evolution, in particular in cases where it is highly invasive, it remains to be determined if such situations are frequent enough to produce a global effect. Here we test this hypothesis using the SymbioCode system, a sample of more than one thousand arthropod species collected in four Polynesian islands (Ramage *et al.*, 2017). We map the previously inferred evolutionary history of *Wolbachia* acquisitions (Bailly-Bechet *et al.*, 2017) on the mtDNA tree, and show that these symbionts affect both the mtDNA polymorphism and substitution rates. While these effects are masked by other sources of variation on a large phylogenetic scale, they become clear once closely related infected and uninfected taxa are compared. These results attest the marked global effect of *Wolbachia* infections on mitochondrial diversity and evolution.

Material and Methods

Dataset

The present analysis is based on the SymbioCode dataset which has been described in details elsewhere (Bailly-Bechet *et al.*, 2017; Ramage *et al.*, 2017) (dx.doi.org/10.5883/DS-SYMC). In brief 10,929 arthropod specimens were collected in four islands of the Society Archipelago, and sorted into morpho-species. DNA barcodes, that is, a 658 bp fragment of the CO1 mitochondrial gene, were obtained by Sanger sequencing from 3,627 specimens, spanning most of the taxonomic and geographic diversity of the initial sample (GenBank ids: KX051578 - KX055204). DNA barcodes clustered into 1,110 species-like groups or Operational Taxonomic Units (OTUs) covering 26 orders, the most species-rich being Diptera (306 OTUs), Lepidoptera (222), Hymenoptera (171), Hemiptera (132) and Coleoptera (106). The presence of *Wolbachia* was ascertained by a double PCR assay (Simões *et al.*, 2011) in 32% of the specimens, and a standard *Wolbachia* marker, the fbpA gene, was directly sequenced from PCR products using Sanger sequencing (Bailly-Bechet *et al.*, 2017). Specimens carrying uncharacterized *Wolbachia* strains (detected by PCR but not successfully sequenced) were excluded from the subsequent analysis, where sequence data was required. To eliminate possible cases of transient infections or artificial contaminations, we also filtered out OTUs carrying a single infected specimen from the present analysis.

A co-phylogenetic analysis based on the tree reconciliation program ALE (Szöllősi *et al.*, 2013a; b) allowed us to map 1,000 plausible scenarios of *Wolbachia* acquisitions on the host mtDNA tree, sampled according to their likelihood (Bailly-Bechet *et al.*, 2017). From these, we estimated the probability that any branch was infected as the proportion of scenarios where this branch was found infected. Notably, this proportion might differ between the start

and the end of a branch (if infection was lost or acquired on this branch); we thus used the average of the two proportions as a measure of the infection probability.

Nucleotide diversity and dN/dS estimations

Within each OTU, the silent nucleotide diversity (π_s) was approximated as the mean of raw genetic distances between all specimens at the 3rd position of codons in the CO1 gene using the R function *nuc.div* (PEGAS, Paradis 2010). The π_{s_inf} value corresponds to infected OTUs (that is, those carrying at least two infected specimens) while the π_{s_un} value corresponds to uninfected OTUs.

Synonymous (dS) and non-synonymous (dN) substitution rates were estimated on each branch of the host tree using a substitution mapping approach. In a first step, we estimated the ω parameter (ω=dN/dS) by maximum likelihood, using a homogeneous model, i.e. assuming a single ω parameter over the entire tree. In a second step, synonymous and non-synonymous substitutions are mapped on the tree to compute branch-specific values of ω. This approach was shown to be more accurate than a full maximum likelihood approach, where one would aim at estimating ω separately for each branch (O'Brien *et al.*, 2009; Romiguier *et al.*, 2012).

Technically, the estimation of ω required to split the maximum likelihood CO1 tree (Bailly-Bechet *et al.*, 2017) in 18 clades of computationally manageable size, each including 170 leaves on average. For each clade, this tree and the corresponding CO1 alignment were used to optimize an homogeneous substitution model by maximum likelihood (Yang & Nielsen, 1998) with the program *bppml* (Bio++ *Maximum Likelihood*) (Dutheil & Boussau, 2008). The model included the following parameters: equilibrium base composition, transition/transversion equilibrium ratio (κ=ts/tv), position specific base compositions at

the root, and a single ω. *MapNH* (from the *TestNH* package) (Dutheil *et al.*, 2012) was then used to count the number of synonymous and non-synonymous substitutions, as well as the number of synonymous and non-synonymous sites, on each branch of the trees. For each branch, the non-synonymous and synonymous substitution rates (dN and dS) are the ratio between the number of substitutions and the number of sites.

To compare ω values of infected and uninfected branches, we pooled estimates from all branches within small clades rather than use single branch estimates, for the two following reasons. First, dN and dS are inaccurately estimated on each specific branch, so that a pool of branches produces more robust estimates. Second, there is some uncertainty in the infection status of each branch, captured in our analysis by the estimated probability of infection. When this probability is not 0 or 1, a specific branch cannot be assigned to the uninfected or infected categories. However, by combining several branches, we can produce pooled estimates of ω for the infected and uninfected lineages by weighting data from every branch using its probability of infection. To combine data from neighbouring branches, we defined clades including specimens distant by no more than 0.2 substitutions per CO1 site (using branch length as a distance measure). This threshold was chosen as a compromise to increase the amount of data for each estimation, without exceedingly increasing the heterogeneity within each pool. Within each of the 536 clades thus identified, we calculated ω along infected and uninfected lineages as follows:

$$\omega_{inf} = \frac{dN_{inf}}{dS_{\text{inf}}} = \frac{\sum_{k=1}^{k=K} dN_k * P_k}{\sum_{k=1}^{k=K} dS_k * P_k}$$

$$\omega_{un} = \frac{dN_{un}}{dS_{un}} = \frac{\sum_{k=1}^{k=K} dN_k * (1 - P_k)}{\sum_{k=1}^{k=K} dS_k * (1 - P_k)}$$

Where P_k denotes the probability that branch k is infected and dN_k and dS_k denote the non-synonymous and synonymous substitution rates on branch k (among K branches). This

calculation ensures that the weight of each branch is proportional to its length and the level of confidence in its infection status. We limited this calculation to clades where ω_{inf} or ω_{un} could be estimated from sufficient data, that is, from branches summing to at least 2% in dS. The ω values along infected and uninfected lineages were thus estimated in 178 and 216 clades, respectively, with 176 clades providing estimates for both categories.

Comparative analysis

Phylogenetic inertia can produce strong but spurious correlations between variables, or on the contrary blur correlations between causally linked variables (Felsenstein, 1985). To test the effect of *Wolbachia* on either nucleotide diversity or substitution patterns, we controlled for this effect using paired comparisons. We used the above-defined 536 clades including specimens distant by no more than 0.2 substitutions per CO1 site. Within each clade including both infected and uninfected individuals, we then calculated the statistic of interest (π or ω) for the infected and uninfected categories. For the polymorphism analysis, these were simply the mean π values of each category of taxa. For the dN/dS analysis, the ω_{inf} and ω_{un} were computed as detailed above. For both analyses, we used Wilcoxon paired signed rank tests to assess differences between the infected and uninfected categories.

Results

Does *Wolbachia* reduce mitochondrial polymorphism?

To assess the effect of *Wolbachia*-induced sweeps, we used sequences of the CO1 gene to compare the silent mitochondrial polymorphism of 134 infected and 241 uninfected arthropod species, collected in French Polynesia as part of the SymbioCode project (Bailly-Bechet *et al.*, 2017; Ramage *et al.*, 2017). Despite a trend in the expected direction, this global comparison did not reveal a significant reduction of polymorphism linked with the presence of *Wolbachia* (Figure 1a; mean π_{s_un}=0.54% (SEM=0.07%); mean π_{s_inf}=0.45% (SEM=0.07%); Wilcoxon rank sum test, W=16947, p=0.38). However, this comparison can be confounded by background variation in mutation rates, census population size, or any other factor affecting polymorphism and varying across arthropod clades. To control for these effects, we used a paired comparison on a subset of the data: 54 infected and 138 uninfected species distributed across 18 small clades (of maximum 20% CO1 divergence) including both categories. This more sensitive approach validates the hypothesis that *Wolbachia* reduces the mitochondrial polymorphism (Figure 1b, Wilcoxon paired test, V=28, p=0.04), with an overall twofold reduction associated with the presence of *Wolbachia* (mean π_{s_un}=1.1%; mean π_{s_inf}=0.51%; mean difference = 0.58%; SE of the mean difference = 0.28%; n=18 clades). Importantly, this effect is visible in clades from several arthropod orders (Figure 1b; π_{s_un} is larger than π_{s_inf} in only 5 clades out of 18), although the signal is necessarily less clear in those harboring a very low polymorphism.

Does *Wolbachia* reduce purifying selection efficiency?

The spread of *Wolbachia* temporarily reduces the mitochondrial effective population size, which produces the above-documented reduction in polymorphism. But are these sweeps frequent and intense enough to also affect substitution patterns, that is, to increase the rate of fixation of non-synonymous mutations, most of which would otherwise be prevented by purifying selection? We tested this hypothesis by comparing ω, i.e. the ratio between non-synonymous and synonymous substitution rates (dN/dS) among infected and uninfected lineages.

To this end, we used the output of a *Wolbachia* / mtDNA cophylogenetic analysis to estimate the probability that *Wolbachia* was present on each branch of the CO1 tree (Bailly-Bechet *et al.*, 2017) and also estimated the number of synonymous and non-synonymous substitutions for each branch. To reduce the uncertainty in our analysis, we did not directly use branch specific estimates, but rather pooled the information from closely related branches, that is, branches belonging to the same clade of maximum 20% CO1 divergence. We further selected the pooled estimates that were based on sufficient total branch length (total dS > 2%). The final dataset thus includes 178 estimates of ω_{inf}, and 216 estimates of ω_{un}, distributed across 218 clades of maximum 20% CO1 divergence.

Regardless of the presence of *Wolbachia*, we found that non-synonymous substitution rates are very low in all lineages, reaching about 1% of the synonymous substitution rate, in line with strong purifying selection acting on the mitochondrial CO1 gene (James *et al.*, 2016). To assess the effect of *Wolbachia* on the efficiency of selection, we first compared the infected and uninfected ω values using a global, non-paired approach (Figure 2a). Although the average infected ω is slightly larger than the average uninfected ω, this non-paired test does not reveal a significant difference (mean ω_{inf}=0.0086, n=178 (SEM=0.0005); mean

ω_{un}=0.0077 (SEM=0.0004), n=216; Wilcoxon rank sum test, W=17530, p=0.13). To better control for the effect of background variations in ω, we selected the 176 clades where both the infected and uninfected ω values could be computed and compared. This paired approach, illustrated in Figure 2b, indicates a significant difference (Wilcoxon paired test, V= 5685, p=0.003), with a 13% increase in ω associated with the presence of *Wolbachia* (mean ω_{inf}=0.008; mean ω_{un}=0.0077 ; mean difference=0.001; SE of the mean difference = 0.0004).

Discussion

Theory suggests that *Wolbachia* infections should reduce the mitochondrial effective population size, while case studies indicate this can occasionally impact the polymorphism (Turelli *et al.*, 1992; Solignac *et al.*, 1994; Ballard *et al.*, 1996; Johnstone & Hurst, 1996b; Keeling *et al.*, 2003; Charlat *et al.*, 2009; Graham & Wilson, 2012; Richardson *et al.*, 2012; Schuler *et al.*, 2016) and possibly the efficacy of purifying selection (Shoemaker *et al.*, 2004). The comparative approach used here suggests these effects are strong enough to globally affect the polymorphism and the molecular evolution of mitochondria. The presence of *Wolbachia* appears to be associated with a twofold reduction in polymorphism, and a 13% increase in the dN/dS ratio. We also note that other factors that are not the focus of this study, such as variation in census population size or mutation rates among different arthropod orders (Allio *et al.*, 2017) introduce substantial variations in mitochondrial diversity and evolution on a large phylogenetic scale, so that the effects of *Wolbachia* are only detected with a paired approach, where we compare closely related infected and uninfected species or branches. This means that *Wolbachia* is one among several forces shaping the mitochondrial polymorphism and substitutions patterns across arthropods.

The observed reduction in polymorphism in *Wolbachia* infected species supports the hypothesis that natural selection acting on the symbiont has produced recent reductions in mitochondrial effective population size in many species. Under this view, several more specific and non-mutually exclusive hypotheses can be distinguished. First, and most obvious, it might be that the recent reduction in Ne was caused by the recent spread of new *Wolbachia* infections, as documented in several case studies (Turelli *et al.*, 1992; Richardson *et al.*, 2012; Schuler *et al.*, 2016). A second possibility is that ancient infections are subject to recurrent selective sweeps, associated with repeated episodes of *Wolbachia* adaptive evolution within its host, which would reduce Ne beyond the initial invasion phase. Finally, a long-term reduction in Ne could also occur if the equilibrium prevalence is low. Indeed, the uninfected part of the population is an evolutionary dead end (that does not contribute to mitochondrial Ne), so that a low equilibrium prevalence can in principle maintain an abnormally low polymorphism in the long run (Johnstone & Hurst, 1996a). Recent estimates of the *Wolbachia* turnover suggest that most infections have been acquired during the last few million years (Bailly-Bechet *et al.*, 2017), a time frame that does not rule out any of the above explanations.

We observed in our dataset that in species where both infected and uninfected lineages coexist, the mitochondrial polymorphism tends to be smaller in the infected portion than in the uninfected portion (π_s=0.35% versus 0.53%). Theory and case studies indicate that once a maternally inherited symbiont has reached its equilibrium prevalence, the reduction in polymorphism also affects the uninfected part of the population, because this part is only maintained through the loss of infection from the infected lineage (Turelli *et al.*, 1992; Solignac *et al.*, 1994; Richardson *et al.*, 2012). Our results indicate that infected species (including those also carrying uninfected specimens) have a lower mitochondrial polymorphism than uninfected ones (0.51% versus 1.1%), suggesting the equilibrium prevalence has been reached in many cases, but also that the infected portion of infected

species harbor an even lower polymorphism (0.35%), suggesting the equilibrium infection prevalence has not yet been reached in a substantial proportion of species. In this context, we also note that species carrying only one infected specimen were removed from the analysis; this allows us to eliminate natural or artificial *Wolbachia* DNA contaminations, but also tends to exclude infected species with low *Wolbachia* prevalence.

While a reduction of polymorphism provides hints on recent selective sweeps, the ω (i.e. dN/dS) ratio integrates all substitutions having occurred over long periods, and can thus reveal long-term variations in Ne. We found a 13% increase in ω associated with the presence of *Wolbachia* on branches of the CO1 tree. We can use previous estimation of the distribution of fitness effects (DFE) of non-synonymous mutations in mitochondria (James *et al.*, 2016) to evaluate the magnitude of change in Ne that would be compatible with our observations. James et al. (2016) estimated the shape of the DFE from over 500 animal species. The shape parameter of this distribution provides a simple relationship between the proportion of non-synonymous mutations becoming effectively neutral (Ohta, 1977; Kimura, 1979; Welch *et al.*, 2008) when Ne is reduced: $x^{-\lambda}=p$ (where x is the factor of change in Ne, λ is the shape parameter of the DFE, and p is the proportion of effectively neutral mutations). From this, we can derive an estimation of x, knowing p and λ: $x=\exp^{-(\log(p)/\lambda)}$. We estimated that omega is 1.13 times larger in *Wolbachia* infected lineages. Assuming the majority of non-synonymous substitutions are deleterious, this means a 1.13 increase in the proportion of effectively neutral mutations. In other words, we estimate p=1.13. James et al (2016) estimate a global λ of 0.44, so that $x=\exp^{-(\log(1.13)/0.44)}=0.76$. Thus, we estimate that Ne is reduced by 24% in lineages where *Wolbachia* is present, which, considering the various sources of uncertainty, is not incompatible with the 50% reduction in Ne estimated from the silent polymorphism data.

The global effect of *Wolbachia* on mitochondrial polymorphism and evolution argues against the view that *within species*, *Wolbachia* might be frequently transmitted between

different maternal lineages, either through occasional paternal transmission (Hoffmann *et al.*, 1990), or horizontally transfer *sensu stricto* (Huigens *et al.*, 2000, 2004). Thus, while non vertical transmission is known to occur and clearly underlies the global distribution of this symbiont (Werren & Windsor, 2000; Engelstädter & Hurst, 2006; Zug *et al.*, 2012; Bailly-Bechet *et al.*, 2017), its rate appears too low to break the genetic linkage between *Wolbachia* and mitochondria *within* species.

Some central features of mitochondrial evolution should be revisited in the light of our findings. Notably, it has been shown that mitochondrial polymorphism is often disconnected from the true population size (Hurst & Jiggins, 2005; Bazin *et al.*, 2006; Galtier *et al.*, 2009). Our results suggest that *Wolbachia* might be part of the explanation, since infected species generally harbour an abnormally low diversity. Evidence is also accumulating that coevolution between mitochondrial and nuclear genes often produces incompatibilities between recently isolated populations, thus contributing to the evolution of reproductive barriers (Burton *et al.*, 2013; Chou & Leu, 2015; Hill, 2016). Specifically, it is hypothesized that mitochondrial properties (high mutation rate, maternal inheritance and lack of recombination) are responsible for the fixation of deleterious or selfish alleles, producing a red-queen-like cytonuclear coevolution. Under this view, mitochondria would represent an Achilles' heel for adaptive evolution, driving compensatory evolution in the nucleus. The increased rate of non-neutral mitochondrial evolution in *Wolbachia* infected lineages might further exacerbate this process.

References

Allio, R., Donega, S., Galtier, N. & Nabholz, B. 2017. Large variation in the ratio of mitochondrial to nuclear mutation rate across animals: implications for genetic diversity and the use of mitochondrial DNA as a molecular marker. *Mol. Biol. Evol.* **msx197**.

Bailly-Bechet, M., Simoes, P., Szöllősi, G., Mialdea, G., Sagot, M.-F. & Charlat, S. 2017. How long does Wolbachia remain on board? *Mol. Biol. Evol.* **34**: 1183–1193.

Ballard, J.W.O., Hatzidakis, J., Karr, T.L. & Kreitmant, M. 1996. Reduced variation in Drosophila simulans mitochondrial DNA. *Genetics* **144**: 1519–1528.

Bazin, E., Glemin, S. & Galtier, N. 2006. Population size does not influence mitochondrial genetic diversity in Animals. *Science (80-.).* **312**: 570–573.

Burton, R.S., Pereira, R.J. & Barreto, F.S. 2013. Cytonuclear Genomic Interactions and Hybrid Breakdown. *Annu. Rev. Ecol. Evol. Syst.* **44**: 281–302.

Charlat, S., Duplouy, A., Hornett, E.A., Dyson, E.A., Davies, N., Roderick, G.K., *et al.* 2009. The joint evolutionary histories of Wolbachia and mitochondria in Hypolimnas bolina. *BMC Evol Biol* **9**: 64.

Chou, J.Y. & Leu, J.Y. 2015. The Red Queen in mitochondria: Cyto-nuclear co-evolution, hybrid breakdown and human disease. *Front. Genet.* **6**: 1–8.

Dutheil, J. & Boussau, B. 2008. Non-homogeneous models of sequence evolution in the Bio++ suite of libraries and programs. *BMC Evol. Biol.* **8**: 255.

Dutheil, J.Y., Galtier, N., Romiguier, J., Douzery, E.J.P., Ranwez, V. & Boussau, B. 2012. Efficient selection of branch-specific models of sequence evolution. *Mol. Biol. Evol.* **29**: 1861–1874.

Engelstädter, J. & Hurst, G.D.D. 2006. The dynamics of parasite incidence across host

species. *Evol. Ecol.* **20**: 603–616.

Felsenstein, J. 1985. Phylogenies and the comparative method. *Am. Nat.* **125**: 1–15.

Galtier, N., Nabholz, B., Glémin, S. & Hurst, G.D.D. 2009. Mitochondrial DNA as a marker of molecular diversity : a reappraisal. *Mol. Ecol.* **18**: 4541–4550.

Graham, R.I. & Wilson, K. 2012. Male-killing Wolbachia and mitochondrial selective sweep in a migratory African insect. 1–11.

Hill, G.E. 2016. Mitonuclear coevolution as the genesis of speciation and the mitochondrial DNA barcode gap. *Ecol. Evol.* **6**: 5831–5842.

Hoffmann, A.A., Turelli, M. & Harshman, L.G. 1990. Factors affecting the distribution of cytoplasmic incompatibility in Drosophila simulans. *Genetics* **126**: 933–948.

Huigens, M.E., de Almeida, R.P., Boons, P.A., Luck, R.F. & Stouthamer, R. 2004. Natural interspecific and intraspecific horizontal transfer of parthenogenesis-inducing Wolbachia in Trichogramma wasps. *Proc Biol Sci* **271**: 509–515.

Huigens, M.E., Luck, R.F., Klaassen, R.H., Maas, M.F., Timmermans, M.J. & Stouthamer, R. 2000. Infectious parthenogenesis. *Nature* **405**: 178–179.

Hurst, G.D.D. & Jiggins, F.M. 2005. Problems with mitochondrial DNA as a marker in population , phylogeographic and phylogenetic studies : the effects of inherited symbionts. *Proc. Natl. Acad. Sci. U. S. A.* **272**: 1525–1534.

James, J.E., Piganeau, G. & Eyre-Walker, A. 2016. The rate of adaptive evolution in animal mitochondria. *Mol. Ecol.* **25**: 67–78.

Johnstone, R.A. & Hurst, G.D.D. 1996a. Maternally inherited male-killing microorganisms may confound interpretation of mtDNA variation in insects. *Biol. J. Linn. Soc.* **53**: 453–470.

Johnstone, R.A. & Hurst, G.D.D. 1996b. Maternally inherited malelkilling microorganisms may confound interpretation of mitochondrial DNA variability. *Biol. J. Linn. Soc.* **58**:

453–470.

Keeling, M.J., Jiggins, F.M. & Read, J.M. 2003. The invasion and coexistence of competing Wolbachia strains. *Heredity (Edinb).* **91**: 382–388.

Kimura, M. 1979. Model of effectively neutral mutations in which selective constraint is incorporated. **76**: 3440–3444.

Martinez, J., Longdon, B., Bauer, S., Chan, Y.S., Miller, W.J., Bourtzis, K., *et al.* 2014. Symbionts Commonly Provide Broad Spectrum Resistance to Viruses in Insects: A Comparative Analysis of Wolbachia Strains. *PLoS Pathog.* **10**: e1004369.

O'Brien, J.D., Minin, V.N. & Suchard, M.A. 2009. Learning to count: Robust estimates for labeled distances between molecular sequences. *Mol. Biol. Evol.* **26**: 801–814.

O'Neill, S.L., Hoffmann, A.A. & Werren, J.H. 1997. *Influential Passengers : Inherited Microorganisms and Arthropod Reproduction*. Oxford University Press, Oxford.

Ohta, T. 1977. Extension to the neutral mutation random drift hypothesis. In: *Molecular evolution and polymorphism: Proceedings of the Second Taniguchi International Symposium on Biophysics* (M. Kimura, ed), pp. 148–167.

Paradis, E. 2010. pegas : an R package for population genetics with an integrated – modular approach. *Bioinformatics* **26**: 419–420.

Ramage, T., Martins-Simoes, P., Mialdea, G., Allemand, R., Duplouy, A.M.R., Rousse, P., *et al.* 2017. A DNA barcode-based survey of terrestrial arthropods in the Society Islands of French Polynesia: host diversity within the SymbioCode project. *Eur. J. Taxon.* **272**: 1–13.

Richardson, M.F., Weinert, L.A., Welch, J.J., Linheiro, R.S., Magwire, M.M., Jiggins, F.M., *et al.* 2012. Population Genomics of the Wolbachia Endosymbiont in Drosophila melanogaster. *PLoS Genet.* **8**: e1003129. Public Library of Science.

Romiguier, J., Figuet, E., Galtier, N., Douzery, E.J.P., Boussau, B., Dutheil, J.Y., *et al.* 2012.

Fast and robust characterization of time-heterogeneous sequence evolutionary processes using substitution mapping. *PLoS One* **7**: e33852.

Schuler, H., Köpler, K., Daxböck-horvath, S., Rasool, B., Krumböck, S., Schwarz, D., *et al.* 2016. The hitchhiker ' s guide to Europe : the infection dynamics of an ongoing Wolbachia invasion and mitochondrial selective sweep in Rhagoletis cerasi. 1595–1609.

Shoemaker, D.D., Dyer, K.A., Ahrens, M., McAbee, K. & Jaenike, J. 2004. Decreased diversity but increased substitution rate in host mtDNA as a consequence of Wolbachia endosymbiont infection. *Genetics* **168**: 2049–2058.

Simões, P.M., Mialdea, G., Reiss, D., Sagot, M.-F. & Charlat, S. 2011. Wolbachia detection: an assessment of standard PCR protocols. *Mol. Ecol. Resour.* **11**: 567–72.

Solignac, M., Vautrin, D. & Rousset, F. 1994. Widespread occurence of the proteobacteria Wolbachia and partial incompatibility in Drosophila melanogaster. *Comptes Rendus l'Academie des Sci. Paris, Série III* **317**: 461–470.

Szöllősi, G.J., Rosikiewicz, W., Boussau, B., Tannier, E. & Daubin, V. 2013a. Efficient exploration of the space of reconciled gene trees. *Syst. Biol.* **62**: 901–12.

Szöllősi, G.J., Tannier, E., Lartillot, N. & Daubin, V. 2013b. Lateral gene transfer from the dead. *Syst. Biol.* **62**: 386–397.

Turelli, M., Hoffmann, A.A. & McKechnie, S.W. 1992. Dynamics of cytoplasmic incompatibility and mtDNA variation in natural Drosophila simulans populations. *Genetics* **132**: 713–723.

Welch, J.J., Eyre-walker, A. & Waxman, D. 2008. Divergence and Polymorphism Under the Nearly Neutral Theory of Molecular Evolution. *J. Mol. Evol.* **67**: 418–426.

Werren, J.H., Baldo, L. & Clark, M.E. 2008. Wolbachia: master manipulators of invertebrate biology. *Nat. Rev. Microbiol.* **6**: 741–51.

Werren, J.H. & Windsor, D.M. 2000. Wolbachia infection frequencies in insects: evidence of a global equilibrium? *Proc Biol Sci* **267**: 1277–1285.

Yang, Z. & Nielsen, R. 1998. Synonymous and Nonsynonymous Rate Variation in Nuclear Genes of Mammals. *J. Mol. Evol.* **46**: 409–418.

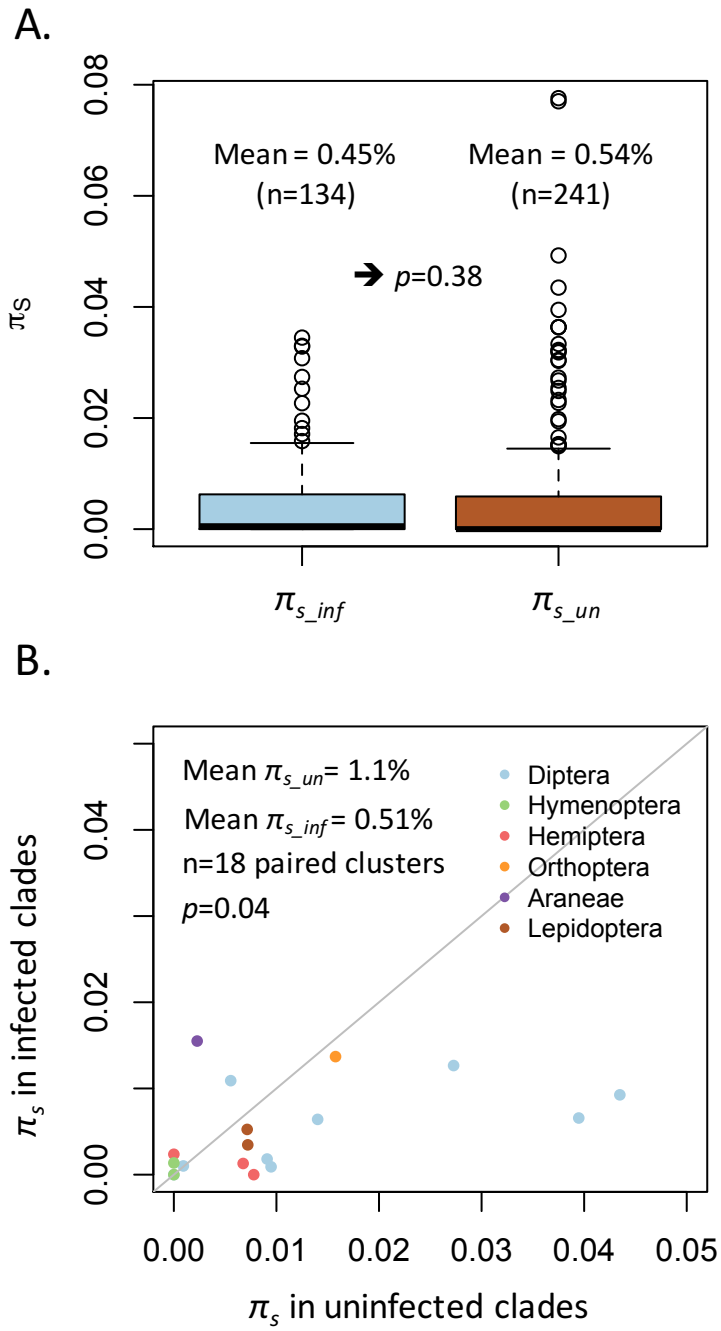
Zug, R., Koehncke, A. & Hammerstein, P. 2012. Epidemiology in evolutionary time : the case of Wolbachia horizontal transmission between arthropod host species. 1–12.

Figures legends

Fig. 1. The effect of *Wolbachia* on silent nucleotidic diversity. A: non paired comparsion between 134 π_{s_inf} and 241 π_{s_un} values (in blue and brown, respectively). B: paired comparison, plotting the mean π_{s_inf} and mean π_{s_un} of 18 clades carrying both categories of species. Two points are on the diagonal, 5 are above, and 11 are below.

Fig. 2. The effect of *Wolbachia* on non-synonymous substitution rates, standardized by synonymous substitution rates (ω=dN/dS). A: non-paired comparisons between 178 ω_{inf} values and 216 ω_{un} values (in blue and brown, respectively). B: paired comparison, plotting ω_{inf} versus ω_{un} across 176 clades where both ω_{inf} or ω_{un} were computed. Two points are on the diagonal, 108 are above, and 66 are below.

Figures



460 **Figure 1**

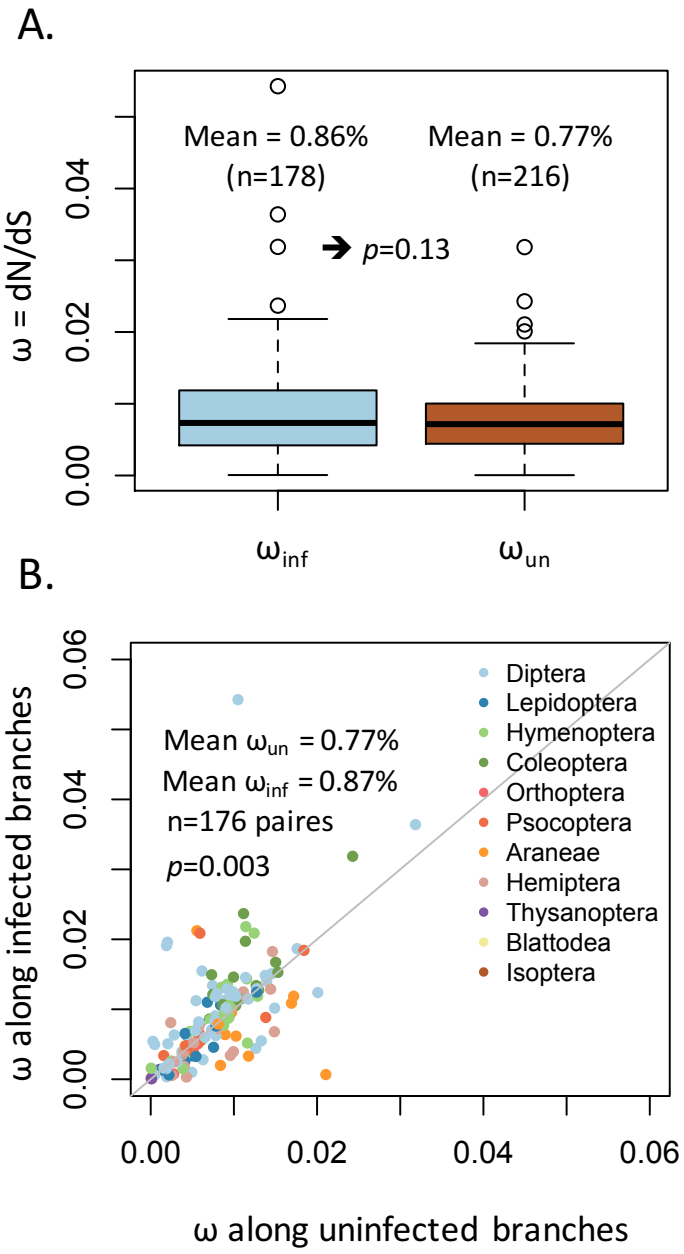

A.
Mean = 0.86%
(n=178)
Mean = 0.77%
(n=216)
p=0.13
ω = dN/dS
0.00
0.02
0.04
ω_inf
ω_un
B.
Mean ω_un = 0.77%
Mean ω_inf = 0.87%
n=176 paires
p=0.003
Diptera
Lepidoptera
Hymenoptera
Coleoptera
Orthoptera
Psocoptera
Araneae
Hemiptera
Thysanoptera
Blattodea
Isoptera
ω along infected branches
ω along uninfected branches
0.00
0.02
0.04
0.06

Figure 2

465